对人类知识和学术的结构性和系统性分析
 作者顾立帆
（未经与我作认真和正式的确认，请不要采用我文章里的任何观点和想法，以免产生任何误解或曲解）
人类的知识和学术，经历了从经验、到数学、到科学的发展。其间不断受到蒙昧主义、诡辩等等的困扰和误导。亚里斯多德谈到了第一性，并建立逻辑学试图清
除诡辩。
但二十世纪的新研究证明了逻辑学不但不具有科学和智力上的第一性和完整性，逻辑学本身也有问题，并无法真正识别和解决很多重要的诡辩问题。亚里斯多德
思维模式和逻辑实证主义均存在问题。而计算机思维模式的问题更加严重。
计算机无法处理高阶逻辑，而只能使用过度简化的逻辑，所以不能处理诡辩，也不具有科学和智力上的第一性和完整性。而诡辩会在司法、医疗、教育、研究等
等方面造成严重问题。从小习惯计算机过度简化的逻辑和流程的学生，会缺乏第一性和完整性的教育训练、缺乏系统性和抽象性的思考能力等等，其智力发展会
受到损害。
至少从 leukotomy 开始，脑和智力研究就陷入错误无法纠正的困境。人工智能的哲学基础和测试理论方法都有严重问题 [1]。拒绝别人用科学实验对
经验结论作证伪，也使诡辩在学术界盛行。诡辩会混淆是非对错，误导科学和文明的发展。而智能技术的新发展，也对 Total Quality
Management 形成了新的挑战。
脑和智力的深入研究，需要对人类的知识、学术、和智力作结构性和系统性研究。神经网络的研究是动物生理层次的，无法真正理解人类特有智力的结构和原理。
而混淆动物性和人的特有智力，用权势压制、迫害学术争论，会导致丛林化和文明瓦解。
本文以动物层次的智力：情绪和经验，和人类独特的智力：语言、哲学、数学、科学，引入智力的结构性和系统性研究，并以此分析亚里斯多德思维模式、逻辑
实证主义、唯物辩证法的问题，分析 Turing Test、leukotomy、AlphaGo Zero、和无人汽车的问题，研究人工智能在哲学基础和
测试理论上的问题，并以思维模式演变的临界点为基础上对人类智力作系统性的分析，最后导出语言处理、科学、和教育发展的新理论。
科学研究基于平等讨论和公正实验。本人设计了一个关于教育、科技、和文明发展的讨论；也设计了严格控制条件、渐进地测试人工智能的科学实验，并以智力
的结构性、系统性分析，和第一性、完整性、复杂性分析，形成智力研究的科学基础。
一、情绪、经验、思维模式
动物主要依靠情绪来影响其他个体。经验也是动物广泛拥有的智力。Neuroscience 主要在视觉、听觉、运动、情绪等等动物层次的研究方面取得进
展。人工神经网络技术，实际上只有经验层次的智力。
人类在语言、哲学、数学、科学等等方面，有独特的智力和思维优势。这些方面的智力原理，超出了 neuroscience 的研究范围。其中很多问题也无法
由神经网络技术解决。所以需要新的研究方法。
由于人类也是动物，也具有动物性的反应模式。这些动物性的反应模式，很多时候比人类特有的思维模式更符合本能。经验和情绪，有动物层次的价值。但人类
也很容易被经验、情绪操纵和误导 [2]。
一个典型的例子是：紧急情况下的战逃反应（Fight-or-flight response）。一进入这个反应模式，人的动物性的情绪和运动反应迅速增强，
但思考能力大大降低。
利用经验和情绪，操纵他人的思维模式、降低别人的思考能力、误导舆论，会增强人的动物性、造成社会的丛林化、甚至使文明崩溃。现在的世界，已经面临这

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible] O.J.Simpson [illegible]

[illegible] Turing Test [illegible]

[illegible]

[illegible] [3][illegible]

[illegible]

[illegible]logical positivism[illegible] positive [illegible] positivism [illegible]

[illegible]

[illegible] positive [illegible]

[illegible]

[illegible]

[illegible] [4][illegible]

[illegible] [5][illegible]

[illegible]Technological Singularity[illegible]

[illegible]Karl Popper[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible] BRAIN Initiative [illegible]

[illegible] [6][illegible]

[illegible] leukotomy[illegible]AlphaGo Zero[illegible]

[illegible]Leukotomy [illegible]

Moniz [illegible]"for his discovery of the therapeutic value of leucotomy in certain psychoses."[illegible] leukotomy [illegible] leukotomy [illegible]

[illegible]leucotomy [illegible]leucotomy [illegible] leucotomy [illegible]

Leukotomy [illegible]

1[illegible] personalities [illegible] mental diseases[illegible]

2[illegible] leukotomy [illegible] leukotomy [illegible]

3[illegible] personality [illegible] intelligence [illegible] Walter Freeman [illegible] personality [illegible] intelligence [illegible] [7][illegible]

无法正确理解 personality 和 intelligence，无法正确区分 personalities 和 mental diseases，无法正确检测对
personality 和 intelligence 的损伤，是由于缺乏正确描述 personality 和 intelligence 的语言。缺乏适当的语言基础，
就无法深入思考和分析。这些问题，目前仍然存在。所以人类的基因改造和基因控制也有很大危险，特别对脑和智力发展。
Leukotomy 获得诺贝尔奖之前，没有作严格控制条件的科学实验和检测，仅仅凭经验性的手术尝试作主观的评估。这种主观评估跟后来发现的灾难性后
果明显不符。对 Leukotomy 有错不纠，也造成现在脑研究项目的问题：SyNAPSE、Human Brain Project、BRAIN
Intiative [8]。
而缺乏严格控制实验条件的主观评估，也正是 Turing Test 的问题之一。Nature 上 AlphaGo Zero 论文中的 superhuman
宣称，也是主观评估。这个 superhuman 是针对 generic 的 human，既缺乏严格控制条件的实验依据，又不接受别人设计的证伪实验。
Leukotomy 的诺贝尔奖有错不纠，和 Nature 上 AlphaGo Zero 论文中的 superhuman 宣称拒绝严格控制条件的实验证
伪，显示 peer review 有严重问题。Peer review 也是主观评估 [9]。使用人工智能技术的无人汽车，也有测试和评估的问题。
这些问题，显示亚里斯多德思维模式和逻辑实证主义有严重问题。下面讨论 AlphaGo Zero 和无人汽车的问题。
四、AlphaGo Zero 的 Superhuman 宣称不是科学结论
Nature 杂志上的 AlphaGo Zero 论文，宣称其取得 superhuman 的 performance。但 superhuman 的概念
是针对 generic 的 human。该论文并没有提供任何能证明 superhuman 的证据。
AlphaGo 系列下过的 game 只是经验测试，不是严格控制条件的科学实验，不能得出科学结论。科学结论只能从严格控制的实验条件推导出来。
AlphaGo Zero 击败 AlphaGo Master 并不是 superhuman 的证据，因为两者都存在人工智能的某些类似缺陷。电脑围棋在
一些 game 中胜过对计算机和人工智能的本质问题尚缺乏了解的棋手，也不是 generic 的 superhuman 的证据。因为人类可以设计科学实
验发现 AlphaGo Zero 的问题，并击败它。
AlphaGo Zero 用算法或数学方法来快速模拟对弈、积累经验、提高棋力。这些模拟对局的数量虽然很大，但跟围棋所有可能走法的数量相比，却微
不足道，远远不够覆盖大部分的可能性。所以 AlphaGo Zero 在绝大部分的情况下，并不知道正确（即最好）的走法，只是靠特征提取和概率性估算
来猜测胜负的可能性。
但传统棋手的棋力，基于经验性智力，在绝大部分情况下也不知道正确走法。水平高的棋手知道比较好的走法，但不知道最好的走法。AlphaGo Zero
经历的模拟对局数，远远超过棋手的对局数。AlphaGo Zero 能总结出的棋力也远远超过人。即使人能抓住 AlphaGo Zero 的错着，但人
的错着更多、错误更严重。很难把优势保持到最后。
所以人类靠经验方法，很难击败 AlphaGo Zero。目前电脑围棋能击败人类棋手，实际上是数学对经验的胜利。但人的智力并不仅仅限于数学和经验。
我设计了 4 维实验空间的科学实验来验证电脑围棋存在大概率的问题，让人类棋手有机会发挥他们在抽象思维和其他方面的优势 [10]。
很多人把深度学习的特征提取，等同于抽象思维。我设计的实验，可以验证两者之间的区别；可以验证对所提取特征的匹配应用，其实只是概率性猜测，并不可
靠。
目前对电脑围棋研究的一个主要问题，是无法正确判断其走法的对错。即使怀疑电脑围棋某些走法是错着，但如果人类棋手最后输了，就难以判断电脑围棋的那
些走法是真正的错着、还是另有胜机的妙着。电脑围棋的特征提取和概率性估算也说明不了问题，因为它也是猜的，并不知道真正的正确走法。
不知道对错就无法确定电脑围棋的真正问题，当然更难以改进。我设计的实验空间，则有方法可以在公平对等的情况下，暴露电脑围棋一些大概率的真正错着和
问题，帮助人类棋手深入研究电脑围棋的问题，帮助学术研究者研究人工智能的系统性问题。

[illegible] AlphaGo Zero [illegible]

[illegible]

[illegible] AlphaGo [illegible] AlphaGo [illegible] [11][illegible]

[illegible] Turing Machine [illegible] AlphaGo [illegible] AlphaGo Zero [illegible] AlphaGo Master[illegible] AlphaGo Zero [illegible] AlphaGo Zero [illegible] [12][illegible]

[illegible]AlphaGo Zero [illegible] Deep Blue [illegible]AlphaGo Zero [illegible] Deep Blue [illegible]Demis Hassabis [illegible] [13][illegible] AlphaGo Zero [illegible] superhuman [illegible]

[illegible] Deepmind [illegible] [14][illegible]

[illegible]

[illegible]

[illegible]:“Go gaming is strictly defined within a very small space. Industrial automations are typically designed in well controlled environments, but not strictly defined. Car driving is regulated, but the environment is not well controlled”[illegible]

[illegible] AlphaGo Zero[illegible]

[illegible] SAE level 5 [illegible] SAE level 4 [illegible] SAE level 4 [illegible] SAE level 4 [illegible]

[illegible] 19X19 [illegible]AlphaGo Zero [illegible]

[illegible] address [illegible]

只是经验测试，不是科学实验。经验测试甚至无法发现波音 737Max 飞机的一些皮毛问题，而使波音公司受到重挫。无人汽车的潜在问题，要严重得多，
也更难用经验测试事先发现。
现在虽然有关于 adversarial testing 和 autoML 等研究，但这些无法系统地研究未来模式的不稳定演变，也无法理解人类智力的结构性以
及人类驾驶汽车时处理意外事件和不稳定演变的能力，甚至无法对自动驾驶的要求写出完整准确可操作的 specification 来作检验。
Deepmind 和 Waymo 的管理者对人类智力的一些关键因素非常无知。本来这种无知，通过对 AlphaGo Zero 的针对性实验，很容易
发现和验证。拒绝科学的证伪实验，不但会造成巨大的资源浪费，还会给未来带来很大的灾难。
人类智力也不仅仅是试错。硅谷高科技风险创业的快速试错文化，有允许 70%的失败、来追求 10%的成功的传统。如果把这种文化用于涉及公众生命安
全的交通管理，会非常危险。而拒绝对人工智能技术用严格控制条件的科学实验来证伪，更使快速试错方法变成毫无科学依据的盲目行为。
等无人汽车大规模应用若干年以后，出现不稳定模式演变、导致大概率严重问题，而无法解决，其灾难性后果远远不是现在波音公司的问题可以相比的。
另外，无人驾驶汽车的测试，完全不同于其他汽车。汽车中的动力（燃料或电动）、机械、电子等设计，早就具备坚实的科学基础，有很强的可控制性。而无人
驾驶的人工智能，没有科学基础，很容易出现急剧的不稳定模式演变，危险性很大。
现在使用的是经验测试，无法解析未来模式的演变。而普通人可能需要很多年跟无人驾驶汽车的大规模互动，才能接触到其临界点，引发重大的模式演变。所以
需要对智力的基础理论作前瞻研究，设计严格控制条件的科学证伪实验来验证和有选择性地使用和控制人工智能技术。否则，人工智能汽车的 Total
Quality Management 根本不可能。
现代制造业的管理模式，从泰勒制和福特制注重生产操作流程、到通用汽车和 Quiz/Whiz Kids 注重目标任务、到 Total Quality
Management 和日本企业注重质量和成本的精细管理。目前的人工智能技术本质上缺乏稳定性和可控制性，不但对很多功能难以作逻辑上的分解，而
且其不稳定性很难控制在子系统内。所以需要新的研究来确保智能技术的 Total Quality Management。
本人不直接研究自动驾驶。而主张对智力作系统性、结构性、和复杂性的基础分析，对电脑围棋等等人工智能技术设计严格控制条件的证伪实验，来研究不稳定
的模式演变，为智力研究建立科学基础。能这样在不引发大规模严重后果的情况下，用基础研究的方法，为自动驾驶技术作前瞻性研究，其实是一种幸运。
本人赞成研究电脑围棋，也赞成研究自动驾驶。只是不赞成过度炒作无人汽车，不赞成把未经严谨的证伪科学实验检测的人工智能技术用于商业化的自动驾驶。
下面分析人工神经网络和预测术的本质，分析基于经验和数学的人工智能和科学的本质不同。
六、预测术、人工神经网络、和计算机技术的局限性
人工神经网络实际上是对已有信息的改写和预测。但科学本身并不是预测，而是在控制条件下对原理的研究。人工神经网络的预测，无法对控制条件深入研究。
人类文明的早期，预测术曾被普遍看重。中国古代有用于预测术的易经：《连山》据说是夏朝的易经，《归藏》据说是商朝的易经，以及周朝的《周易》。
商代夏，周代商，模式演变大致是稳定的。旧易经不对，就换新易经。到春秋战国的混乱取代周朝的实际统治时、和秦国结束战国混乱时，模式演变是不稳定的。
没有人再能写出流行的新易经。预测术虽仍继续存在，但已远不如夏、商、周时那样重要了。
《道德经》中的“道可道，非常道。名可名，非常名。”实际上是对预测术的否定。即使是现代的科学之道，也需要不断发展。
到欧洲文艺复兴后期，由于天文学的发展，占星预测术又开始流行。
第谷的天文理论体系是错的，却得到了很多资助。开普勒的天文理论体系是对的，却死于贫困。开普勒的天文学研究，在其生前并不受人重视，不得不靠占星预
测术来争取收入，维持他的天文学研究。

[illegible]

[illegible]

[illegible]

[illegible] [15][illegible]

[illegible]

[illegible]

[illegible] Turing Machine [illegible] Universal approximation theorem [illegible]

[illegible] Technological Singularity[illegible]

[illegible]

[illegible]

1949 [illegible] leukotomy [illegible]

Turing Test [illegible] AI[illegible]A Modern Approach [illegible] Wind Tunnel approach [illegible] Nature [illegible] AlphaGo Zero [illegible] superhuman [illegible]Technological Singularity[illegible]In Math We Trust[illegible]“[illegible]”[illegible] [16][illegible] [17][illegible]

[illegible]

[illegible] [18][illegible]

[illegible]Occam's Razor[illegible]

Occam's Razor [illegible] Occam's Razor [illegible]

[illegible]
Occam's Razor [illegible] Leukotomy [illegible]

[illegible]

[illegible]

[illegible]
Gu Test：A Progressive Measurement Of Generic Artificial Intelligence [illegible]

[illegible]

[illegible]

[illegible]

[illegible] [19]。

[illegible] [20]。

[illegible]

[illegible]

[illegible]

[illegible]

[illegible] commonsense [illegible]

[illegible]

[illegible]

九、数学和科学的演变及其重要临界点
从毕达哥拉斯到伽利略，是人类文明走上可持续发展的关键。
毕达哥拉斯的抽象化思维，创立了数学。他对算术和数的定义和体系，按现代抽象代数的概念也是成立的，但不适用于几何和其他一些系统。毕达哥拉斯算术系
统本身在科学上是错的，但在数学上又是可以成立的、有价值的、可以继续发展的。所以数学不同于科学，也不能简单用数学来验证科学上的对错。
希帕索斯的发现，本来也是对数学和毕达哥拉斯体系的发展。但希帕索斯事件之后，毕达哥拉斯学派本身就没有太多成就了。其他群体则能继续他们的工作，后
来居上。
欧几里德几何创建了重要而又可以完备化的数学理论体系，并带动了阿基米德等等一大批泛希腊化时代的学术研究者。
毕达哥拉斯和欧几里德两千年之后，有牛顿和莱布尼茨创建微积分，罗巴切夫斯基等人推翻欧几里德几何的唯一性，建立非欧几何 [21]。
朗道实际上把自己当作第四流的科学家：牛顿是第一流，爱因斯坦是第二流，创建量子物理的才俊们是第三流，和朗道同时代的美国物理学家理查德·費曼也是第
三流。也许朗道是自谦，但他肯定无法和爱因斯坦相比。而俄罗斯历史上是出过门捷列夫的。
不知朗道该如何评价伽利略、达尔文。
按照我对学术成就的评价，创建重要学科、和重要学科中的理论体系，是第一流的工作；在重要学科中提出革命性理论体系，是第二流的工作；对主要理论体系
的重要延伸或修补、或开创较小的学术分支，可列入第三流的成就。第一、第二流的学术成就非常罕见，第三流的工作已经非常了不起了。这个标准跟朗道的评
价高度吻合。
物理学中，伽利略和牛顿的贡献，跟数学中的毕达哥拉斯和欧几里德相仿，属第一流。相对论和量子力学是重要学科的革命性理论体系，跟数学中的微积分和非
欧几何相仿。所以，爱因斯坦的贡献是第二流的，而创建量子力学有许多人的贡献，他们的贡献只能算第三流的。量子电动力学只是小分支，所以理查德·費曼的
工作也是第三流的。
牛顿和莱布尼茨在数学中的贡献是第二流的。但非欧几何的创建，由多人逐步完成，有许多人的贡献。罗巴切夫斯基和鲍耶对推翻欧几里德几何的唯一性最重要，
他们和希帕索斯（发现无理数）、亚里斯多德（逻辑学）、帕斯卡（概率论）、哥德尔（哥德尔定理）相仿，仅略逊于牛顿和莱布尼茨。
拉格日朗（数学分析的成熟化）、图灵（图灵机作为计算模型）的贡献是第三流的；数学中的高斯、和物理学中的朗道，能不能挤进第三流，我不知道；
祖冲之的贡献，未必能超过高斯。但有观点认为《九章算术》首次引入负数 [22]。此说如果确实，《九章算术》的数学成就，也许并不比希帕索斯和亚里斯
多德逊色很多。这可能跟中国古代的阴阳玄学哲学性观念有关 [23]。
数学是非常古老的学科，现在看起来很重要的一些成就，跟数学历史长河中的很多大师比起来，不一定更重要。十七、十八世纪的数学，和二十世纪前半期的物
理学类似，已经十分兴旺，但基础是更早的人奠定的。
达尔文跟毕达哥拉斯和伽利略相仿。孟德尔不如欧几里德和牛顿，因为生物学整体理论在不久之后还需要几次重要的综合才能形成。
波义耳和门捷列夫在化学中的作用，有点类似于物理学中的伽利略和牛顿。但门捷列夫和波义耳不但时间相隔较久，而且中间有很多人的重要贡献。所以，他们
的学术贡献，也不如伽利略和牛顿。
这里的评价，全部按照学术贡献的大小。不是评估个人智力，也不是评价人的尊卑。而是评估其工作对人类整体智力发展的影响。人各有特长，个人智力也不一
定都表现出来，很多情况下是不可比较的。而学术贡献，是已完成的工作，可以客观分析。
整理历史上学术贡献的大小，一方面是为了研究人类智力的演变和其中重要的临界点，另一方面也是研究能促进未来发展的各种因素。现在把智力研究转化成科
学的关键，在于形成准确描述智力的语言概念、哲学基础、和数学形式。

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible] Chinese room [illegible]

[illegible]

[illegible] dataset [illegible] SQuAD[illegible]CoQA[illegible]QuAC[illegible] GLUE [illegible] dataset [illegible]Chinese room [illegible]

NLVR^2[illegible]Natural Language for Visual Reasoning for Real[illegible] testset [illegible]GLUE [illegible] generic [illegible]

[illegible] Testsets [illegible]AI: A Modern Approach[illegible]

[illegible] guideline [illegible] judgement[illegible]

[illegible] Chinese room [illegible]

[illegible]

The Third Wave [illegible]

[illegible] AlphaGo [illegible] self-driving car [illegible]

The Third Wave [illegible]

[illegible]

[illegible] Total Quality Management [illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible] Leukotomy [illegible]AI: A Modern Approach[illegible]

[illegible]

[illegible]“[illegible]”[illegible] [24][illegible]

[illegible]

了对科学思想方法的训练。而现在不成熟的教育 AI 技术，会使这种问题更加严重。
没有这种思维训练，当然就分析不出 AlphaGo Zero、语言处理和诡辩、和 AI: A Modern Approach 中的问题，也无法从科学哲
学的高度抽象概括出亚里斯多德思维模式和计算机思维模式的系统性和结构性问题，甚至会使未来的科学研究误入歧途。
大学主流人工智能教科书 AI: A Modern Approach 的哲学基础和测试理论都有严重错误，AI 本身就有严重问题，怎么能靠 AI 来自我进
化、纠正编教科书的大学教授的错误？
思维模式的问题，儒家教育造成的万马齐喑悲剧就是一个典型例子。洋务运动从知识层面作改进，在科学和教育方面的提高也有限，根本不能解决问题。
对高等教育来说，如果靠亚里斯多德的信徒编辑知识点、给出规划，哥白尼、开普勒、伽利略都会被封杀，怎么能发展现代科技？
儒家和亚里斯多德的信徒都是人，至少还具有人的一般智力和变通适应能力。而计算机和人工智能技术，使用过度简化的逻辑，有 Chinese room 等
等的问题，不能真正理解语言、哲学、数学、科学等等人类学术的发展和构造。如果让学生从小就服从计算机过度简化的逻辑和流程，人的特质会退化，智力发
展会受严重损害，后果更严重。
现在有人工智能教育方面的人士，把对称思维、创造性思维、分类讨论思维、类比思维等等当作思维模式。这些其实只是提供更多思维可能性的思考方法，不包
含决定结论对错的 judgement，也不是思维模式。
而思维模式，会决定结论的对错。很多思维模式，如计算机思维模式、亚里斯多德思维模式，会导致错误的结论。所以需要研究很多思维模式的问题，也需要研
究科学是怎么避免和纠正很多思维模式的问题。
计算机和人工智能技术，既给教育带来很大帮助和机遇，也带来很大挑战。需要对语言处理、思维模式等等有关人工智能的问题，用科学方法认真研究。在科学
验证的基础上，谨慎地设计教育中的人机耦合。
特别需要谨慎的是，中小学教育实际上是红海市场。整体教育规模很难有大幅度增长，前些年在计算机和网络教育上也已经投入很多资金、形成了很多用户
[25]。
而科学新学科的建立和发展，至少可能给高等教育带来局部的增长。所以，人工智能教科书的问题，知识、学术、智力的结构系统理论，如何严格控制条件来设
计证伪实验的科学基础，这些研究对未来的发展非常关键。
研究性大学，并不等同于精英大学。科学需要平等讨论和公正实验。平民因此可以靠科学上的努力，弥补一些自己的弱势。而开创科学新领域，探索新的教育和
研究方式，更需要多样化的教育，理解不同的教育模式和需要。
十三、不同的教育模式
1989 年，新加坡举行了一场辩论赛，辩题是：“儒家思想可不可以抵御西方歪风”。辩论虽然是反方获胜，但评委最后的结论却是辩论双方都没有把问题分
析清楚。当然评委自己也无法分析出对错。
辩论赛和 AlphaGo 系列的围棋比赛一样，虽然有胜负，却不能分辨学术上的对错 [26]。辩论赛和 Socratic 方法，也有本质不同。这作为智力
发展的问题，需要进一步研究。
如果把这场辩论赛，和 2015 年的 Bohunt Chinese School 实验、及 BBC 的 Are Our Kids Tough Enough
? 的相关系列电视节目对照起来看，就有意思了。
实验中的中国教师一再提到英国学生落后中国学生三年。这个说法并不准确。有数据说上海学生在 2012 年 PISA 评估中最好，领先各国平均学生三年。
但 2015 年和 2018 年 PISA 用中国发达的东部两市两省代表中国大陆后，中国大陆发达地区学生的表现，不但落后于新加坡，也落后于台湾等好几
个国家和地区，跟上海学生单独得分差距非常大。上海、新加坡的领先，是否跟这两城市的户籍、入境管理和其他因素有关，也需要研究。
所以，用数据作依据，很容易产生误导。这是数据方法需要警惕的问题。Bohunt Chinese School 实验，只是经验性的实验，不是科学实验。

Bohunt 校长本人也谈到实验结果不能说明问题 [27]。跟人有关时，特别需要注意经验实验和科学实验的区别。
整体上，新加坡、香港、澳门、台湾、中国大陆的华人聚居区，确实在 PISA 评估中名列前茅。这里有个文化原因。华人在七世纪开始系统化、制度化的考试，
远远领先其他国家，也远远早于十七世纪的科学革命。但华人在考试上的悠久文化，并没有催生出科学革命。
Bohunt 实验中的中国教师，把 Confucianism（孔儒）当作哲学基础，不但没有说服力，反而在英国学生中造成了很大的困惑，并在一定程度
上妨害了他们的教学。本来他们的教学可以更加成功的。
事实上正是孔儒文化直接导致了中国近代在科学上的严重落后。孔儒文化，讲究传承。但跟现代教育的内容，却有严重脱节。这可以用 Bohunt 实验中的中
国教师教授的几门学术课程：数学、科学、英文、中文来作分析。
英文出现和成形，跟孔儒完全没有关系。科学的起源和成形，跟孔儒也基本没有关系。数学的起源和成形，和孔儒也关系甚微。中文倒跟孔儒关系较大。但孔子
之前中文的文言文已经比较成熟；以后的发展是百家争鸣，连普通人都有很多贡献；二十世纪初的新文化运动更是通过对孔儒的激烈批判才发展起现代中文。
所以，对英国学生谈文化传承，讲莎士比亚、科学革命、工业革命，才跟课程真正有关系，才能让他们理解在发展中的真实自我。而更深的渊源，需要追溯到毕
达哥拉斯、希帕索斯、苏格拉底、欧几里德等等。现在人工智能在哲学基础和测试理论上的问题，以及对智力的结构性和系统性分析，都需要在深度上理解这样
的文化传承。
孔儒是主张尊卑有序的。而那几位中国教师反复讲的 discipline 和 competition，其实是比较现代的概念，跟孔儒理念并不相同。事实上正是
尊卑观念妨害了他们的教学。
Discipline 跟尊卑无关，只跟人在不同场合的角色有关。在 Bohunt 实验中，很多英国学生对英国教师很尊重，但对中国教师经常嘲弄、甚至愚弄。
那些英国教师对学生有管理上的权势；而中国教师来自相对不发达的国家，很容易被认为地位卑下。跟英国学生讲 Socratic 方法、平等讨论，才能从根本
上消除他们的嘲弄、愚弄态度。
那几位中国教师反复讲 competition，强调要变得更强。这其实是对达尔文学说的误解。达尔文已经认识到是适者生存，而不是强者生存。科学家和教
师，跟权势和财富的拥有者比起来，其实比较弱势，很容易在骨子里受鄙夷。
宋朝灭亡后，郑思肖在《心史》中就提到“一官、二吏、三僧、四道、五医、六工、七猎、八娼、九儒、十丐。”做不了官的教师是儒，在元朝地位比妓女都低。
而元朝正是中国历史上最强大的朝代。英国伦敦在 2012 年刚刚竖起成吉思汗的雕像。
所以，不宜让学生误以为学业优秀就是强者。追求学术上的进步是为了帮助人类更好地适应，是为信仰而服务。社会则应该尊重这种服务。作为信仰和服务，需
要认真分辨学术上的对错。教育上是十年树人，更要对学生负责、分清对错，不能颠倒是非。
那几位中国教师，虽然对讲授的专业比较纯熟，但对专业以外的很多问题有些片面、狭隘。对自己的处境、社会中实际存在的尊卑次序、孔儒尊卑有序观念的真
正后果，也未必了解 [28]。
但英国在教育上确实面临问题。十九世纪，英帝国和清帝国，是世界上管理人口最多的两大国家。英帝国虽然战胜了清帝国，但在达尔文之后，其学术研究跟其
他发达国家相比，其实在走下坡路 [29]。到了李约瑟、鲁桂珍时代，思想更是有肤浅化的趋势。对不同思维模式及其教育模式，并不真正了解。中国和其他
非西方的思维模式和教育模式，并非一无是处，但不是李约瑟和孔儒所讲述的那种优势。Leukotomy 的错误无法纠正，更给未来脑和智力的研究蒙上阴
影。
理解不同文化和不同教育模式，需要对智力和文化作结构性、系统性的深度分析，需要新的哲学。现在，美国教育也有类似于英国教育的问题，甚至更严重。
Technological Singularity 和主流教科书 AI: A Modern Approach 中的问题，就是这种迹象。
十四、有关教育、科技、和文明发展的讨论

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]——[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible] Technological Singularity [illegible] [30][illegible]

[illegible]

[illegible]

[illegible] Karl Popper [illegible]

[illegible]

BRAIN Initiative [illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible] The Development of Liberal Arts and Sciences [illegible]

~~~~~~~~~~~~~~~~~

[1] [illegible] AI: A Modern Approach [illegible]“Aristotle… was the first to formulate a precise set of laws governing the rational part of the mind.”(On page 5)[illegible]

[illegible] Wind Tunnel approach [illegible]

[illegible]

[2] [illegible]

[illegible] Technological Singularity[illegible]AlphaGo Zero [illegible] superhuman [illegible]In Math We Trust [16][illegible]“[illegible]” [illegible]

[illegible]

[illegible]

[illegible]

[illegible]“Read my lips: no new taxes”[illegible]

[illegible]

[illegible]“[illegible]”[illegible]

[illegible]

[illegible] Quiz/Whiz Kids [illegible] Pentagon Papers [illegible]MBA [illegible]

[illegible]
~~~~~~~~~~~~~~~~~

[illegible]

[3] [illegible]: [illegible] metaphysics [illegible]

[4] [illegible]

O.J.Simpson [illegible]

[illegible] [20][illegible]O.J.Simpson [illegible]O.J.Simpson [illegible]

[5] [illegible]

[illegible]

[illegible]

[6] [illegible]

[illegible]

[7] [illegible] Leucotomy in England and Wales, 1942-1954[illegible] 9284 [illegible]41[illegible]28[illegible]25[illegible]2[illegible]4[illegible]

[illegible] personality [illegible] intelligence[illegible] 25[illegible] personality [illegible] intelligence [illegible] clinical condition [illegible]41[illegible] 28[illegible] clinical condition [illegible]personality [illegible] intelligence [illegible]

[illegible]leucotomy [illegible]

[illegible] Renato M.E. Sabbatini [illegible]Even lobotomy's preponents admitted that only one third of the operated patients would improve, while one-third remained the same, and one-third got worst [illegible] Leucotomy in England and Wales, 1942-1954 [illegible]http://www.cerebromente.org.br/n02/historia/lobotomy.htm[illegible]

[illegible] one third would improve [illegible] one-third remained the same[illegible] clinical condition[illegible] personality [illegible] intelligence [illegible]

[8] SyNAPSE [illegible] ACM Gordon Bell prize[illegible]Dharmendra Modha [illegible] thank you note[illegible]

Henry Markram [illegible] SyNAPSE [illegible] annoucement [illegible] mass deception of the public[illegible] SyNAPSE [illegible]Henry Markram [illegible]"It is not impossible to build a human brain and we can do it in 10 years."[illegible] Human Brain project[illegible]

而按奥巴马总统的要求，美国也开始了被 NIH Director 称为 moonshot 的 BRAIN Initative 项目。该项目以基因技术和基于计
算机大数据技术的 dynamic brain activity map 为基础，而无视计算机技术和生命、智力的本质不同，大数据技术也无法理解智力的本质、
无法准确理解对智力的损害。而目前的 neurosciences 方法对生理和动物层次的智力研究有帮助，如视觉、听觉、运动、情绪等等；对人的特有智力
现象难以理解。计算机技术在这方面不但没有帮助，反而会有误导。事实上这种基于计算机技术的巨型脑研究项目确实压制了对人脑和计算机本质不同的研究，
误导了脑和人工智能的研究。
脑研究是科学研究，跟 moonshot 的工程项目有本质不同。后面要谈到的钱学森教授的火箭设计，才跟 moonshot 有些关系。而钱学森教授研究
农业展望和人体特异功能，需要用科学方法，跟钱教授的本行工程设计和数学有本质不同。
钱学森教授和 NIH Director 都不理解工程和科学的区别，这是教育有失败之处。
平等讨论、公正实验的科学方法，关系到文明的长远未来，比一时的权势、利益争夺更加重要。
[9] 对学术论文的审核，只代表审核者的观点，不是学术上的定论。所以，学术上的苏格拉底方法很重要。要给被审核者和其他人反驳的机会。
对科学论文的审核，也跟数学论文不同。数学论文只能靠 peer review。科学论文中的结论，可以用新的证伪实验来推翻。其错误，从论文本身可能看
不出。所以，对科学论文的 peer review，并不能保证其正确性。要让别人有机会设计新实验来证伪。
对科学论文的审核，可以发现推理问题，实验数据是否真实等等。但推理问题只能引起注意，有可能是出于理解有误。所以下结论要经过对话、反驳、和实验证
伪。AlphaGo Zero 论文中的 superhuman 宣称，是针对 generic 的 human，并没有相应的证据。本人却可以设计严格控制
条件的实验来证伪此宣称。但 AlphaGo Zero 方面至今拒绝对话和实验。
[10] 我选择围棋作实验工具，是因为可以用围棋验证人工智能和计算机的一些基础科学理论，跟军事战略和战争艺术无关。围棋是和平的竞技，高度抽象而
又直观。既古老又现代，能吸引亚洲以外的电脑研究。已经成为跨文化研究、交流的平台。
第一轮实验后，开发者可能修改软件。所以有些实验证据只能在第一轮实验中收集。第一轮实验需要跟最强版本电脑围棋作较大规模的实验，以收集足够的实验
证据。
我对 AlphaGo 系列的实验不是白盒测试，不涉及知识产权。如果 Google 愿意公正地检验，可以仿照美国国会图书馆先例，先向国会图书馆递交
AlphaGo 柯洁版、AlphaGo Zero、和最新 AlphaGo 版本所需要的所有硬件、软件和文档，能捐款给图书馆用于管理和维护则更好。
以后在第三方场地实验时，可以向国会图书馆核对软件、硬件和文档等版本。Human level artificial intelligence 的检验，是人类
发展史上的里程碑。如果无人能对 AlphaGo 系列证伪，国会图书馆就有继续保留其版本的价值。其他人有实验想法，可以先做。如果其他人都失败了，就
没有道理拒绝我的实验方案。
这也为美国政府了解人工智能的最新进展提供了渠道，并为以后的人工智能测试建立了一个范式。一旦相关实验设施建立好并熟悉了这种流程，以后重要的人工
智能测试费用会大大降低。
[11] 因此实验涉及到某些敏感因素。目前我设计的实验方案，只能发现一部分问题，不能发现全部问题。所以需要进一步的研究。而我的实验，可以提供进
一步研究所需要的证据。所以要防止实验前系统被针对性地改动，破坏了进一步研究所需要的证据。
科学要求通过不断证伪来验证、改进结论。而 Demis Hassabis 公开宣布 AlphaGo Zero 不再参加比赛，拒绝了进一步的验证。事实上
AlphaGo Zero 从来没有严格控制条件下的实验证据。所以如果 Deepmind 拒绝验证，会破坏科学的方法和基础，损害学术道德。
再考虑到 Deepmind 的 ethics board 仅一个星期就被解散，Deepmind 在道德上言行是否一致有很大疑问。我对证据被破坏的担心
并不多余。所以，Google 向国会图书馆递交 AlphaGo 系列的软、硬件和文档后，我才能递交我的实验方案。
[12] 目前能见到的 AlphaGo Zero 棋谱中，对 AlphaGo Master 的二十局是 AlphaGo Zero 最后版（或稳定版），有
很不寻常现象：AlphaGo Master 执黑的十盘棋中，布局至 16 手完全一样。而 AlphaGo Zero 执黑的十盘棋中，布局至 18 手完

全一样，AlphaGo Zero 输棋的 14、16 局，至 45 手完全一样。所以有如下问题：
1）在 Nature Magazime 上发表两篇 AlphaGo 系列文章后，Deepmind 有没有在这二十局棋里把 AlphaGo Zero
和 AlphaGo Master 对局的情况，全面准确地告诉公众？如果答案是，则：
2) 公布的这二十局棋谱中，所出现的布局高度类同，是由于 AlphaGo Zero 接近围棋最优解，还是陷入 local trap？如果认为是前者，需
要用严格控制条件的严谨实验来验证。
[13] 如在 The Guadian 访谈中他谈到 a meta-solution to any problem：
"Demis Hassabis ... is deadly serious when he tells me he is on a mission to 'solve
intelligence, and then use that to solve everything else'.
......
'One way of thinking of AGI is as a process that will automatically convert
unstructured information into actionable knowledge. What we’re working on is
potentially a meta-solution to any problem.'",
from https://www.theguardian.com/technology/2016/feb/16/demis-hassabis-artificial-intelligence-deepmind-alphago。
[14] 许峰雄在 Cracking Go 文中更早提出。但他的依据很多来自开发 Deep Blue 的经验，对围棋不一定成立。我的理论分析，不依赖经
验，反而能成立。我的理论分析还可以进一步延伸：AlphaGo 系列虽然赢棋，但只要不是最优解，就可能走的是无理棋；可以设计严格控制条件的严谨实
验来检验 AlphaGo 系列是否有通用智力潜力，是否超人。
[15] 由三体问题的研究发现，行星的真正轨道极为复杂，传统数学都难以完整表达或逼近。所以，下文提到：Universal approximation
theorem 的名称，也很容易误导；而可以抽象成 Turing Machine 及其各种变种的计算机技术，也有本质的局限性。
[16] 货币是重要的权势。虚拟币的成败，跟背后的权势较量有关，跟数学的关系反而不大。被卷进权势的较量，也很危险。
In God We Trust 是美国的格言。这里提到的 God 是全能神。现代科学可以构造多个具有哲学完整性且无法证伪的科学理论。这些无法证伪的
科学理论，使全能神的存在，也无法被证伪。但人不是全能神，人无法完整理解全能神，所以说 Trust。
数学是人的一种智力理论，不是全能神。数学的局限性，早已被证明。所以，In Math We Trust 并不正确，是迷信。In Math We
Trust 在精英专业研究人士中的流传，暴露了科学哲学教育的缺乏和现代教育的问题。
[17] 登月项目是工程。需要定义明确的实施目标。而脑研究是基础科学研究，需要通过科学实验的证伪来不断修改假设。所以，工程和基础科学研究的方法
完全不同。出现这样的混淆，是教育和研究严重失败所造成的无知。
[18] 在政治宣传、商业宣传中，诡辩和语义的混淆比比皆是，难以清除。所以，科学研究和司法，需要跟政治宣传和商业宣传严格区分，保持应有的独立性。
[19] 数学上的欧几里德几何和非欧几何，是作为有效理论并存的。数学可以作为科学的工具，但本身并不是科学。事实上，非欧几何的出现，就明确宣告欧
几里德几何不是科学，而仅仅是数学，数学也不具有哲学完整性。哥德尔定理更说明数学不能判断科学上的正确性。而科学上的多种替代性理论中只有一种是正
确的，所以科学有完整性。
[20] 所以，达尔文的理论，按其性质的不同，可以分成两部分：
一部分是演化的规律，这是科学理论，可以用实验检验和证伪。虽然充分地检验，需要很长时间。这部分争议较少。所揭示的生物本质和规律，奠定了区别于博

物学的生物科学基础。
另一部分是物种的起源。这和宇宙大爆炸理论一样，实际上是对历史的逆向推导，是历史研究。虽然可以用到很多科学方法，但历史研究本质上不是科学，只能
提供某种程度的证据，不能用严格控制条件的实验验证或证伪。当然，不是科学并不等于就是错的，只是说这种理论达不到科学理论的标准。但历史研究不能用
实验验证或证伪，确实也造成对一些年代较远的事情的结论经常出现变化甚至反复。
[21] 高斯不但没有公布他自己的非欧几何想法，也没有公布 1819 年就交给他的 Ferdinand Schweikart 的类似想法。所以高斯对非
欧几何，不但无功，反而有过。原因就在于高斯太注重坚实的技巧性，不了解尚未成熟的思想体系的价值，也不重视苏格拉底方法对学术发展的作用。高斯不但
缺乏学术眼光，也达不到罗巴切夫斯基的思想境界和层次。
罗巴切夫斯基和鲍耶推翻欧几里德几何的唯一性，已是 1830 年前后了；他们在非欧几何上贡献最大。黎曼几何在物理相对论研究中，有价值。但黎曼在非欧
几何中，只是作延伸性工作，贡献次于罗巴切夫斯基和鲍耶。贝尔特拉米为非欧几何带来新的语义和完整性，也很重要。
由于被高斯埋没，Ferdinand Schweikart 可能没有对非欧几何的发展产生多少实际影响。我也不知道如何评价他的工作。
[22] 《九章算术·方程》：“正负术曰：同名相除，异名相益，正无入负之，负无入正之。其异名相除，同名相益，正无入正之，负无入负之。”
也有观点认为，《九章算术》之前，中国已经出现负数概念。
[23] 中国古代百家争鸣时期，就有阴阳家，《道德经》和解释《易经》的《十翼》也都论及阴阳。阴阳概念的真正起源，有待考证。但《易经》体系的阴阳
并不导出负数，而是：老阳数九，老阴数六，等等。
阴阳概念，不但可以很容易导出正负数，也为现代正反物质的理论提供了一个很自然的通俗理解语言和思维模式。但能相互抵消的正负数，是在阴阳概念基础上
的飞跃。现代正反量子湮没的理论，又是对可抵消的正负数概念的飞跃，因为量子湮没是伴随有能量形态的转化的。从中正好可以看到从玄学、到数学、到科学
的思维模式的演变和发展。
但阴阳和玄学不是科学哲学语言，不是具有完整性的严格意义上的哲学。最多只能算语言层次的准哲学或哲学性思想，进一步发展的潜力有限。在形成正负数以
后，中国古代数学也没有在这个方向有多少后继发展。所以，即使准哲学或哲学性思想能帮助形成一些新的学术思想，准哲学或哲学性思想本身也需要不断发展，
需要新的语言概念。
中国古代没有像希帕索斯那样对无理数的准确发现，对类似无理数的概念没多少抵触，但也不理解其价值。这里有哲学和思维模式的不同，受此影响智力的发展
轨迹也不同。
张首晟教授生前提到过阴阳、正负、和量子湮没。在此注明一下，作为对他的纪念。本人很多年没有参加投资创业活动，以前也没有看到张教授的这个观点。他
去世后搜寻他的信息，才发现他有类似观点。也许科学的追寻者本来就可以心灵相通。写上述内容，是为了整理中国古代学术思想。写作时仅通过
wikipedia 查证正负数概念出自中国。
不过本人的观点跟张教授也有所不同。本人着重指出从不导出负数的阴阳、到相互抵消的正负数、到伴随能量转化的量子湮没，是从玄学、到数学、到科学的不
断发展，其哲学性基础有质的变化。这里包含着思维模式演变的临界点。而张教授等一些学者认为准粒子和基本粒子没有本质区别的观点，是物理学界的争论，
本人不发表看法。
准粒子跟 emergent phenomena 有关，是凝聚态物理中的重要概念。在物理学中，数学推导不能作为科学结论。但物理学中有海森堡测不准
原理。如果临近测量的极限，会有很多争议，有些甚至无法解决。但离测量极限尚有一定距离时，仍有可能通过精化测量的尺度来区分。
[24] 奥林匹克数学竞赛和朗道势垒也是种应试，只是难度较高。奥林匹克数学竞赛在科学性和理论深度、广度、高度上也远远不及朗道势垒。只是一味追求
难度，反而会使教育丢失更重要的目标。
最近用计算机和人工智能技术改进中小学教育的努力，重新重视了教育的基本面、和对基础知识的解析和积累。但计算机和人工智能技术有局限性，也需要深入
研究。

[25]

mainframe personal computer smartphone smartphone

IT

100 100 60

[26]

[illegible]

[27] “[illegible]·[illegible] BBC [illegible]”[illegible]http://shanghai.xinmin.cn/xmsq/2016/04/18/29861595.html[illegible]

[illegible]

[28] [illegible]“[illegible]”[illegible]·[illegible]“[illegible]”[illegible]

[illegible]

[illegible]“[illegible]”[illegible]

[illegible]“[illegible]”[illegible]“[illegible]”[illegible]

[29] [illegible]

[illegible]

[30] [illegible] Nature [illegible] AlphaGo Zero [illegible] superhuman [illegible]

[illegible]

[illegible]